Title: Immediate and distracted imitation in Second-Language speech : unreleased plosives in English

Author: Arkadiusz Rojczyk, Andrzej Porzuczek, Marcin Bergier

Citation style: Rojczyk Arkadiusz, Porzuczek Andrzej, Bergier Marcin. (2013). Immediate and distracted imitation in Second-Language speech : unreleased plosives in English. "Research in Language" (2013, iss. 1, s. 3-18), doi 10.2478/v10015-012-0007-7

IMMEDIATE AND DISTRACTED IMITATION IN SECOND-LANGUAGE SPEECH: UNRELEASED PLOSIVES IN ENGLISH

ARKADIUSZ ROJCZYK
University of Silesia, Poland
arkadiusz.rojczyk@us.edu.pl

ANDRZEJ PORZUCZEK
University of Silesia, Poland
ampj24@wp.pl

MARCIN BERGIER
University of Silesia, Poland
emberg@o2.pl

Abstract
The paper investigates immediate and distracted imitation in second-language speech using unreleased plosives. Unreleased plosives are fairly frequently found in English sequences of two stops. Polish, on the other hand, is characterised by a significant rate of releases in such sequences. This cross-linguistic difference served as material to look into how and to what extent non-native properties of sounds can be produced in immediate and distracted imitation. Thirteen native speakers of Polish first read and then imitated sequences of words with two stops straddling the word boundary. Stimuli for imitation had no release of the first stop. The results revealed that (1) a non-native feature such as the lack of the release burst can be imitated; (2) distracting imitation impedes imitative performance; (3) the type of a sequence interacts with the magnitude of an imitative effect

Keywords: imitation, plosives, unreleased, distraction.

1. Introduction

Human beings have an inborn capacity to reproduce the actions and intentions of others (Hauser 1996; Honorof et al. 2011; Nagell et al. 1993; Whiten and Custance 1996). This imitative tendency starts immediately after birth (Meltzoff and Moore 1999) - for instance, twelve-week old infants already imitate ambient vocalic sounds (Kuhl and Meltzoff 1996) - and appears to reach its climax between two to five years of age (Horner and Whiten 2005). Those early imitative reactions are logically linked with language acquisition processes which encourage children to acquire language from their caretakers and peer group (Chambers 1992; Babel 2012). The automatic imitative behaviour observed in humans appears to have a neurophysiological basis in the architecture of mirror neurons which make up an action-observation matching system (Rizzolatti and Craighero 2004; Rizzolatti et al. 2001; Schwartz et al. 2012). It is suggested that the human mirror-neuron system creates parity between the speaker and the listener, which is a prerequisite for successful imitation (Arbib 2005; Gentilucci and Corballis 2006; Rizzolatti and Arbib 1998). It is achieved by activation in brain areas responsible for planning and production of speech during auditory or visual perception of speech (Pekkola et al. 2006; Pulvermuller et al. 2006; Skipper et al. 2007; Wilson and Iacoboni 2006; Wilson et al. 2004).

The speech-imitative behaviour persists into adulthood and serves many sociolinguistic functions. For example, adults relatively easily acquire features of a dialect in the new surrounding (Delvaux and Soquet 2007; Evans and Iverson 2007; Munro et al. 1999; Trudgill 1986). Such imitation seems to be driven by the need to assimilate with a new positively-evaluated social group, however Bourhis and Giles (1997) also observed a dialect divergence conditioned by a negative affective attitude towards a particular dialect group. Talkers also exhibit imitative tendencies in various communicative interactions to express similarity (Shepard et al. 2001), attraction (Byrne 1971), to gain approval (Street and Giles 1982), or to increase one's intelligibility (Triandis and Triandis 1960).

2. Imitation in speech

Phonetic imitation, also referred to as phonetic convergence or phonetic accommodation, is the process in which a talker takes on acoustic characteristics of their interacting partner (Babel 2012). A variety of phonetic features have been reported to undergo imitative convergence, such as accent, speaking rate, intensity, pitch, variation of frequency bands, long-term average spectra, frequency of pauses, and utterance length (Giles et al. 1991; Goldinger 1997; Gregory 1990; Gregory and Webster 1996; Gregory et al. 1993, 1997, 2001; Namy et al. 2002; Natale 1975; Pardo et al. 2012). Other studies have concentrated on VOT as a temporal parameter that undergoes assimilation as a result of exposure to the model talker. Shockley et al. (2004) demonstrated significant VOT imitation of voiceless plosives in words with artificially extended VOT values. More recently, Nielsen (2011) showed that not only are longer VOTs imitated by talkers but also that this re-modelled feature can be generalized to other plosives. Moreover, imitation in this study was found to be selective, in that it did not occur for reduced VOT and depended on the frequency of tested lexical items. Significant imitation of the model talker has also been reported for vowels, as expressed by formant frequencies of individual productions (Babel 2010; Evans and Iverson 2007; Pardo 2010; Pardo et al. 2010, 2012). Here again, the degree of convergence was modulated by selectivity, in that only some vowels in different linguistic items were imitated. Finally, Honorof et al. (2011) reported convergence of articulatory gestures in imitated allophonic instances /l/, expressed as the distance between F2 and F1.

All the observed imitative tendencies in speech are captured by non-abstract theories of linguistic representations. In this view, fine-grained linguistic and non-linguistic phonetic features available in the speech signal are preserved in perception to make a set of exemplars that forms a perceptual category (Hintzman 1986; Nosofsky 1986). This is the main assumption of exemplar-based models of speech perception (Coleman 2002; Johnson 1997; Pierrehumbert 2006). Although those models do not rule out completely the possibility of modularity in formation of categories, they predict that mental phonological representations of words encode both allophonic variability and speaker-specific information. Following this reasoning, imitation of speech emerges as a natural process in which the listener perceives and reproduces fine-grained phonetic features provided by the model talker. Even if such features differ from the listener's canonical representations, they are not filtered out or discarded, but rather they are successfully delivered in imitation. Such accommodation from perception to production is considered to be both automatic (Gentilucci and Bernardis 2007, but see Nielsen 2011 and Mitterer and Ernestus 2008 for selectivity) and quick (Fowler et al. 2003; Honorof 2011).

Imitation is an undisputed factor in acquisition of second-language speech. Successful production of non-native sound categories should logically arise from effective imitation of patterns absent in one's native language. Previous research in our lab has shown that acoustic features defined by cross-linguistic differences between Polish and English can be imitated to a significant degree by learners when shadowing after the model talker. Rojczyk (2012a) had Polish learners of English imitate the English low-front vowel /æ/ in a rapid shadowing task. This vowel is reported to be difficult to acquire for Poles: it is equally likely to be assimilated by two Polish neighbouring vowels /e/ and /a/. Productions in two tasks were compared: reading of words with the target vowel presented as a list with a view to establishing a baseline condition and imitations of the same words delivered binaurally. F1 and F2 were measured for all productions and the Euclidean distance to the model vowel frequencies was calculated to express the degree of convergence. The results showed that the learners significantly converged their productions of the target vowels with the model. It was taken as evidence that imitation can override the influence of native categories in production of new sound categories. In another study, Rojczyk (2012b) used VOT as another Polish-English typological difference that emerges in Polish pronunciation of English. While English /p, t, k/ are characterised by long-lag VOT values, Polish /p, t, k/ use short-lag VOT values. As mentioned earlier, this difference surfaces in Polish pronunciation of English as observable underaspiration of English voiceless stops. In this study, production in three tasks was compared: (1) reading of English words with /p, t, k/ word-finally as a baseline condition; (2) immediate imitation of those words pronounced with native-like long VOTs; (3) distracted imitation in which the imitators were required to read a digit presented on the screen after hearing a model word and prior to imitation. The results revealed significant increase in VOTs in immediate imitation and intermediate values for distracted imitation. These results were interpreted to indicate that immediate imitation may bypass the influence of native articulatory habits and that distraction in imitation results in incomplete recovery of native phonetic patterns.

3. Release burst in stop sequences in English and Polish

English and Polish differ considerably in the frequency of release bursts both word-finally and when preceding another stop. Many textbooks on English phonetics observe that English stops tend to be unreleased when followed by another stop or affricate (e.g.

Abercrombie 1967; Gimson 2001; Jones 1956; Ladefoged 1975; Roach 2000). Experimental research has supported this observation, however its magnitude is not so great as may be expected. Crystal and House (1988a) found 59% of English stops without the release burst in all sentence positions. Randolph (1989 reported in Byrd 1993) reported that in word-final position English stops are mostly unreleased. Byrd (1993) analysed data from the TIMIT database and found 40.3% of releases in stops. Davidson (2010) investigated spontaneous speech from the National Public Radio and found the frequency of the unreleased stops in pre-stop and pre-pausal position between 50% and 60%. The actual frequency of unreleased variants depends on many factors. Bilabial stops tend to be more often unreleased, followed by alveolars and velars (Byrd 1993; Crystal and House 1988b). Voiceless stops have a stronger tendency to include and acoustically measurable release burst than voiced stops (Crystal and House 1988a, b, but see Byrd 1993). Finally, women have been reported to release stops more often than men (Byrd 1992, 1993).

Polish stops are generally described as invariably released except when they precede another homorganic stop (Dukiewicz and Sawicka 1995; Jassem 1974; Kopczyński 1977; Wierzchowska 1980). Rojczyk (2008) studied experimentally the putative tendency to unrelease Polish stops in same-place clusters. Stops were matched across a word boundary in two-word phrases and sentences. The results revealed that in the case of an intervening word boundary stops were released more than 50% of the time in homorganic clusters. The actual context significantly influenced the frequency of release bursts. Stops inserted in short two-word phrases were more frequently released than stops in sentences. All this leads to the conclusion that Polish differs from English in the tendency to unrelease stops and that this will have consequences on Polish pronunciation of English. Indeed, observations by experienced teachers of English pronunciation indicate that Polish learners have noticeably more frequent and stronger releases in English stop sequences than English native speakers and that this contributes to the perception of their speech as non-native. Textbooks on English pronunciation tailored for Polish learners include exercises in this area (Bałutowa 1974; Mańkowska et al. 2009; Sobkowiak 2001). Although controlling for the lack of release is initially difficult, appropriate phonetic training and instructing can yield positive effects on ultimate performance (Bergier 2010).

4. The current study

The current study investigates the degree of imitation of English unreleased stops by Polish learners. It consists of three tasks: (1) reading of phrases presented as a list to establish a baseline frequency of releases in the studied group; (2) immediate imitation of the unreleased sequences provided by the model talker; (3) distracted imitation of the unreleased sequences in which listeners are required to read a digit after hearing a model stimulus and prior to imitation. Accordingly, the research questions are formulated as follows:

1. Is the lack of the release burst imitated in immediate imitation calculated as a significant decrease in the frequency of releases compared to list-reading?

2. Does distraction in imitation impede the performance compared to immediate imitation or does it block imitative behaviours altogether as compared to list-reading?
3. Does the type of a sequence - same place of articulation vs. different place of articulation - interact with the magnitude of imitation?

Previous research showed that delaying imitation reduces its degree (Goldinger 1998). In this study, we decided to use distraction, which we suggest poses a greater challenge on listeners. While simple delaying extends the time interval between auditory input and articulatory production, it does not interfere with articulatory planning because no other cognitively taxing processes are included. On the other hand, distraction, in which the response is not only delayed but subjects also engaged in reading digits, provides both an increase in cognitive processing and articulatory resetting between hearing and imitating target stimuli. We therefore assumed that a greater challenge to imitators would yield more reliable results on how long-lasting perceptual traces of the lack of release were in the studied group.

Gender was not included as an independent variable largely due to the fact that some of the used statistical procedures used were non-parametric for nominal variables and did not allow inclusion of more than one independent variable. We predict, however, that the investigation of how gender interacts with imitation of unreleased plosives would yield interesting results. In earlier studies women were reported to converge to a larger extent to the model talker compared to men (Namy et al. 2002; Pardo 2006).

4.1 Participants

Thirteen native speakers of Polish (eight females and five males) were included in the study. They ranged in age from 20 to 21. All participants were students at the Institute of English, University of Silesia. None of them had had any prior phonetic training concerning unreleased stops in English. They did not have any reported speech or hearing disorders.

4.2 Materials

The stimuli used in the experiment were nine two-word noun phrases in which stop sequences were matched across the word boundary (Table 1).

ca**p P**at	tha**t p**an	bla**ck p**ack
ta**p t**ap	tha**t t**ap	bla**ck t**ap
la**p c**at	tha**t c**at	bla**ck c**ap

Table 1. Stimulus phrases used in the experiment. Bolding indicates combinations of stops straddling word boundaries

All nouns were preceded by either a verb, adjective or determiner and had a natural focus stress on the second word. The intonation pattern was uniform with a H* H*L-L%

contour. Only voiceless stops were selected for two reasons. First, voiceless plosion is more conspicuous in a spectrographic display because it is not attenuated by concomitant periodicity. Second, voiceless plosives have been previously reported to be more frequently released (Crystal and House 1988a, b). Considering a greater tendency to release voiceless plosives rather than voiced plosives, it was assumed that using voiceless plosives would be a more challenging task on the participants and thus a more sensitive metric of the occurrence of imitative behaviour. The stimuli reflected all possible combinations of places of articulation: bilabial, alveolar and velar (3x3=9) and contained only one, low front vowel /æ/.

All stimuli were recorded by the second author, a qualified phonetician, using the recording specification described below. The stimuli were created by saving individual sound files in a computer. Next, spectrographic analysis in Praat (Boersma 2001) was used to inspect target sequences of two stops. No release bursts were detected in the preceding stops. The duration of each stop sequence was durationally normalized to range from 180 ms to 190 ms. Finally, all stimuli were peak normalized to 70 dB SPL.

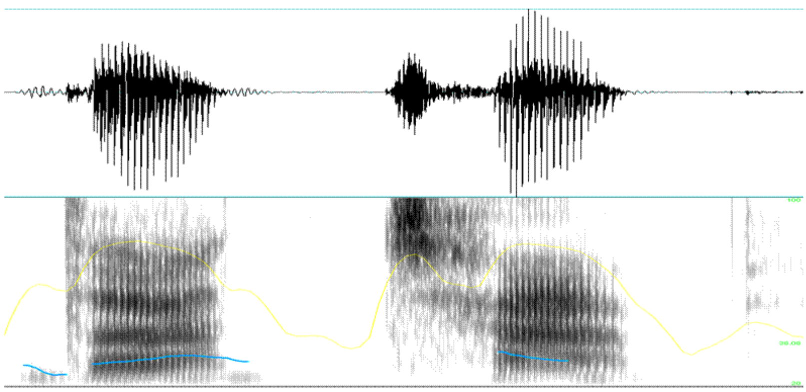

Figure 1. Waveform and spectrogram of the phrase *that tap*. No detectable release burst of the first stop in a sequence.

4.3 Procedure and recording

The experiment took place in the Acoustic-Phonetic Laboratory at the Institute of English, University of Silesia. As described earlier, data were collected from three blocked tasks. The first task was reading the list of phrases presented orthographically to establish a baseline frequency of release bursts in the studied material. The words were flashed sequentially on a monitor screen in 54-point black font in the middle of the white screen. Seven foil phrases were randomly dispersed among target phrases to distract the participants' attention from the object of the study. The second block was immediate shadowing after the model talker in which participants were instructed that upon hearing the model pronunciation they were to immediately repeat it. The orthographic representations of the phrases were also sequentially flashed during imitation. The approximate interval between complete imitation and the onset of the next phrase was 1 sec. The third block was distracted imitation. The participants were instructed that they would hear the model pronunciation, next they would read a digit flashed in the centre of the screen, and finally that they were to imitate the phrase. The interval between playing the model voice, displaying a digit, and flashing an orthographic representation of the imitated phrase was also approximately 1 sec. Tasks 2 and 3 were counterbalanced by participants to avoid a carry-over effect from one task to another.

The recordings were made in a sound-proof booth with a monitor screen located in front of a participant. The signal was captured with a headset dynamic microphone Sennheiser HMD 26, preamplified with USBPre2 (Sound Devices) into .wav format with the sampling rate 48 kHz, 24-bit quantization. The model voice was delivered binaurally through high-quality headphones built in the headset at a comfortable listening level.

4.4 Measurements

All measurements were made using waveform and spectrogram displays available in Praat (Boersma 2001). Prior to any quantifications of data, it was necessary to define the acoustic criteria for classification of measured tokens as released or unreleased. Introductory analyses revealed a number of cases with visible weak energy spikes in the spectrogram but which, at the same time, gave no auditory impressions. Henderson and Repp (1982) listed five stages of the unreleased-released continuum: (1) unreleased; (2) silently released; (3) inaudibly released; (4) weakly released; (5) strongly released. As a result, we decided to classify our tokens as released when they belonged to stages (4) and (5) from Henderson and Repp (1982). In other words, stops were classified as released when they had auditorily detectable burst and it was manifested as the sudden rise of energy visible as acoustic transients in waveform and spectrogram. Other tokens were classified as unreleased.

It is interesting to note that release bursts measured in the current data were characterised by significant variability. While the force of the burst appeared to be, to a large extent, individual, some task effects were observed, especially in distracted imitation. Some sequences had a very long compression phase of the first stop followed by relatively long release, sometimes even exceeding 100 ms. This may show articulatory attempts to hold the compression and proceed to another stop, which however were not successful and ended in a strong release of pent-up air. It is, however, for future experiments to see how detailed acoustic properties of release may reveal imitative behaviour.

Measurements were divided into two main types for statistical analyses. Nominal measurements classified tokens as either released or unreleased. Ratio measurements identified the time duration of the burst expressed in ms. It was assumed that measuring the duration of bursts might be a more sensitive metric of whether imitation occurred or not. Duration of the burst was defined as the time interval between the onset of the rise of energy following a silent period of compression to its offset indicated by a complete drop of energy signalling compression for the next stop. A total number of measured tokens was 351 (13 participants x 9 sequences x 3 tasks).

4.5. Results and analysis

Two types of statistical tests were run to calculate both nominal and ratio data. For nominal observation, i.e. a stop may be released or unreleased, Cochran Q test was applied. It is an alternative to one-way within-subject ANOVA when the dependent variable is dichotomous. For duration measurements in ms a two-way mixed ANOVA 2x3 was designed with 2 levels of a between-subject variable (place of articulation: homorganic / heterorganic) and 3 levels of within-subject variable (task: list reading / immediate imitation / distracted imitation).

Figure 2 shows the overall proportion of release bursts in all three tasks as well as the proportions broken down into homorganic and heterorganic clusters.

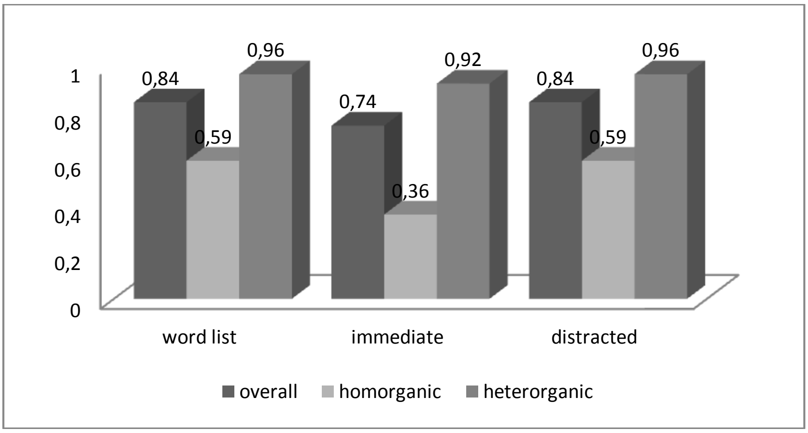

Figure 2. The proportion of release bursts in all three tasks: overall, homorganic clusters and heterorganic clusters.

For an overall number of release bursts the Cochran Q test revealed significant influence of task on releasing a stop [x^2 (2) = 10.67, p < .01]. This effect was achieved by reduction of release bursts in immediate imitation (74%) compared to word list (84%) and distracted imitation (84%). Breaking down data into homorganic and heterorganic clusters showed different magnitude of contribution to the main effect of task. For homorganic clusters there was again a significant effect of task on releasing a stop [x^2 (2) = 9, p < .05]. The number of release bursts decreased in immediate imitation (26%) relative to word list (59%) and distracted imitation (59%). No statistically significant effect of task was found for heterorganic clusters [x^2 (2) = 2, *ns*]. Summing up the analysis of the frequency of release busts, the following observations may be formulated. Immediate imitation reduced the number of release bursts relative to the baseline articulatory habits. This reduction was mostly contributed to by homorganic sequences. When imitation was distracted it did not lead to the reduction of release bursts.

The analysis of duration of bursts in ms was predicted to be a more sensitive measure of imitative behaviour, in that it would be able to register slight reductions of plosion force which were disregarded in a dichotomous released / unreleased measure. Figure 3 shows the overall mean durations of bursts in ms in all three tasks as well as durations broken down into homorganic and heterorganic clusters.

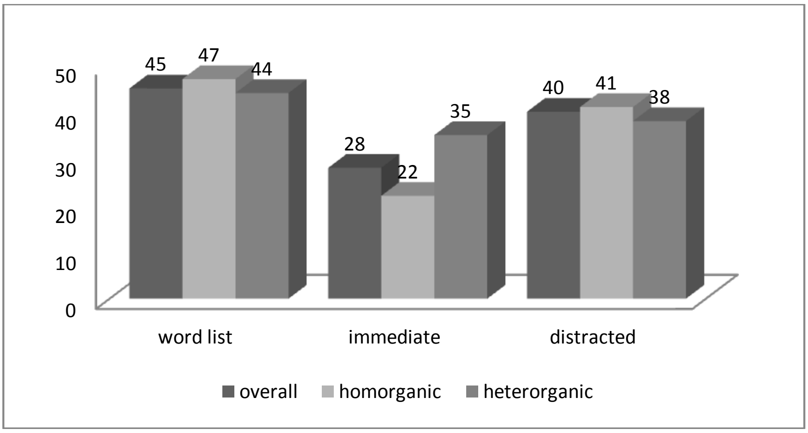

Figure 3. Mean durations of release bursts in all three tasks: overall, homorganic clusters and heterorganic clusters.

The main effect of task on duration of the release bursts was highly significant [F (2, 230) = 15.86, p < .001]. Post hoc Bonferroni tests revealed that this effect was mainly achieved by significant reduction of durations in immediate imitation (28 ms) compared to baseline list-reading (45 ms) (p < .001) and distracted imitation (40 ms) (p < .05). When imitation was distracted, durations of release burst tended to be lower than in a baseline list-reading task, however not significantly (both p > .05).

The task x cluster type interaction was significant [F (2, 230) = 4.38, p < .05], indicating that the effect of task on durations varied in magnitude depending on whether the cluster was homorganic or heterorganic. Post hoc Bonferroni tests revealed that significant reduction of burst durations observed in immediate imitation was mostly contributed to by homorganic clusters in which durations of bursts dropped from 47 ms to 22 ms (p < .001). Although heterorganic clusters also demonstrated reduction in immediate imitation (35 ms) compared to list-reading (44 ms), it was not statistically significant (p > .05). All comparisons between immediate imitation and distracted imitation for either homorganic or heterorganic clusters were non-significant. The same lack of significance was also found for comparisons between baseline list-reading and distracted imitation, which indicates that distraction in imitation resets articulatory patterns to their default status.

5. General discussion

The current study investigated how two types of imitation - immediate and distracted - modify the pronunciation of FL learners. Unreleased plosives in two-stop sequences in English were chosen, because English and Polish differ in the frequency of releases in such sequences. Polish learners of English were exposed to the auditory model that produced unreleased stops in two imitation tasks: immediate imitation in which shadowing commenced immediately after the auditory input and distracted imitation in which participants were instructed to read a digit after the auditory input and prior to shadowing. Two types of measurements were used, nominal released / unreleased and durational in which the durations of bursts were expressed in ms. The results were expected to show if, and to what extent, unreleased stops can be imitated compared to baseline list-reading.

The analysed data allow us to answer questions formulated before the experiment.

1. Is the lack of the release burst imitated in immediate imitation calculated as a significant decrease in the frequency of releases compared to list-reading?

The answer is positive. Both nominal and durational measures revealed that the participants reacted to the auditory input and modified their productions to converge with the model talker. The frequency of bursts and their durations significantly decreased in immediate imitation compared to baseline list-reading. It demonstrates that the lack of release can be imitated even by talkers whose native language releases stops in stop sequences. Future studies should look more closely into individual variability because, as shown in the current data, participants differed in their initial tendency to release and in how they reacted to the auditory model in imitation. Some participants had a high initial release frequency and did not observably reduce it in imitation. Other participants showed a similar initial high release frequency but reduced it as a result of auditory exposure. Yet others had a relatively low release frequency initially and either reduced it or not in imitation. It is for future research to investigate more thoroughly the extent of individual variability both in a default release pattern and in the release pattern in imitation and to seek explanation for those idiosyncrasies.

2. Does distraction in imitation impede the performance compared to immediate imitation or does it block imitative behaviours altogether as compared to list-reading?

The answer is positive. Current results showed that distracting participants by asking them to read digits after the auditory exposure and prior to imitation reduced significantly imitation effect. Although this is evident in calculations of statistical significance, it is worth noting that the imitation effect was not absent altogether. Both the frequency of bursts and their durations were lower for distracted imitation relative to baseline list-reading. It points to some weak remnants of auditory traces despite distraction, as evidenced by participants' productions. As mentioned earlier, some productions in distracted imitation were characterised by long compression phases and sudden strong and long releases. It may be interpreted to mean that the participants attempted to imitate the lack of release by extending durationally the hold phase, but were finally unsuccessful, which resulted in strong releases of compressed air. In order to verify this tendency, future studies may make use of more sensitive acoustic metrics.

Another point that merits discussion is the very nature of distracted imitation compared to delayed imitation. While delayed imitation extends the time interval between auditory exposure and production, distraction provides two additional parameters: cognitive taxing and articulatory resetting. Cognitive taxing is caused by the need to perceive and recognise the digit to plan articulatory commands for its production. Articulatory resetting is a product of a new plan for articulation. This is not the case for delayed imitation in which articulatory resetting does not occur because participants are inactive between the exposure and production. This difference does not seem to be satisfactorily explained in the imitation literature. While it seems safe to assume that distraction will cause more disturbance to imitation than delaying, the actual magnitude of this inhibition is not clear. Future studies should incorporate the distinction between delaying and distraction as a variable to disentangle their effect on imitation.

3. Does the type of a sequence - same place of articulation vs. different place of articulation - interact with the magnitude of imitation?

The answer is positive. Homorganic clusters, unlike heterorganic clusters, had a significant contribution to the observed effect of imitation. We ascribe it to the fact that, as discussed earlier, stops in homorganic clusters in Polish can be optionally unreleased. If homorganic unreleased stops are classified as allophones of released stops in Polish, it is clear evidence that what is most readily imitated is the allophonic variant which occurs in participants' native language. Previous research has demonstrated that allophonic experience from learners' native language improves both perception of non-native contrasts as well as their learning. (Best and Strange 1992; Halle et al. 1999; Jamieson and Moroson 1986; Jenkins and Yeni-Komshian 1995; Kondaurova and Francis 2008; McAllister et al. 2002; McClaskey et al. 1983; Pisoni et al. 1982; Pruitt et al. 2006). The current results suggest that such experience also contributes to imitation. It is not surprising considering the fact that successful imitation is an important factor in developing perception of non-native sounds and in successful acquisition of their production. A special status of native language allophonic variants reported here appears to confirm their special status in acquisition of second-language speech in general.

References

Abercrombie, D. 1967. *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.

Arbib, M. A. 2005. From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioural and Brain Sciences* 28: 105-167.

Babel, M. 2010. Dialect convergence and divergence in New Zealand English. *Language in Society* 39: 437-456.

Babel, M. 2012. Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics* 40: 177-189.

Bałutowa, B. 1974. *Wymowa Angielska dla Wszystkich*. Warszawa: Wiedza Powszechna.

Bergier, M. 2010. *The occurrence of unreleased oral stops in English voice agreeing plosive clusters straddling word boundaries. Production experiments with Polish advanced learners of English.* M. A. Thesis. University of Silesia.

Best, C. T. and W. Strange. 1992. Effects of language-specific phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics* 20: 305-330.

Boersma, P. 2001. Praat, a system for doing phonetics by computer. *Glot International* 10: 341-345.

Bourhis, R. Y. and H. Giles. 1977. The language of intergroup distinctiveness. In H. Giles (ed.) *Language, Ethnicity, and Intergroup Relations*. London: Academic Press: 119-136.

Byrd, D. 1992. Preliminary results on speaker-dependent variation in the TIMIT database. *Journal of the Acoustical Society of America* 92: 593-596.

Byrd, D. 1993. 54,000 American stops. *UCLA Working Papers in Phonetics* 83: 1-19.

Byrne, D. 1971. *The Attraction Paradigm*. New York: Academic Press.

Chambers, J. 1992. Dialect acquisition. *Language* 68: 673-705.

Coleman, J. 2002. Phonetic representations in the mental lexicon. In J. Durand and B. Laks (eds) *Phonetics, Phonology and Cognition*. Oxford: Oxford University Press: 96-130.

Crystal, T. H. and A. S. House. 1988a. The duration of American-English stop consonants. *Journal of Phonetics* 16: 285-294.

Crystal, T. H. and A. S. House. 1988b Segmental durations in connected-speech signals: Current results. *Journal of the Acoustical Society of America* 83: 1553-1573.

Davidson, L. 2010. Variation in stop releases in American English spontaneous speech. *Journal of the Acoustical Society of America* 128: 2458.

Delvaux, V. and A. Soquet. 2007. The influence of ambient speech on adult speech productions through unintentional imitation. *Phonetica* 64: 145-173.

Dukiewicz, L. and I. Sawicka. 1995. Fonetyka i fonologia. In H. Wróbel (ed.) *Gramatyka i Fonologia Języka Polskiego*. Kraków: Instytut Języka Polskiego PAN.

Evans, B. G. and P. Iverson. 2007. Plasticity in vowel perception and production: A study of accent change in young adults. *Journal of the Acoustical Society of America* 121: 3814-3826.

Fowler, C. A., J. M. Brown, L. Sabadini and J. Weihing. 2003. Rapid access to speech gestures in perception: Evidence from choice and simple response time tasks. *Journal of Memory and Language* 49: 396-413.

Gentilucci, B. R. Bernardis. 2007. Imitation during phoneme production. *Neuropsychologia* 45: 608-615.

Gentilucci, B. and M. C. Corballis. 2006. From manual gesture to speech: A gradual transition. *Neuroscience and Behavioral Reviews* 30: 949-960.

Giles, H., J. Coupland and N. Coupland. 1991. *Contexts of Accommodation: Developments in Applied Sociolinguistics*. Cambridge: Cambridge University Press.

Gimson, A. C. 2001. *Gimson's Pronunciation of English*. 6th edition. Rev. by A. Cruttenden. London: Arnold.

Goldinger, S. D. 1997. Perception and production in an episodic lexicon. In K. Johnson and J. W. Mullennix (eds) *Talker Variability in Speech Processing*. San Diego: Academic Press: 33-66.

Goldinger, S. D. 1998. Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105: 251-279.

Gregory, S. W. 1990. Analysis of fundamental frequency reveals covariation in interview partners' speech. *Journal of Nonverbal Behavior* 14: 237-251.

Gregory, S. W., K. Dagan and S. Webster. 1997. Evaluating the relation between vocal accommodation in conversational partners' fundamental frequencies to perceptions of communication quality. *Journal of Nonverbal Behavior* 21: 23-43.

Gregory, S. W., B. E. Green, R. M. Carrothers and K. A. Dagan. 2001. Verifying the primacy of voice fundamental frequency in social status accommodation. *Language Communication* 21: 37-60.

Gregory, S. W. and S. Webster. 1996. A nonverbal signal in voices of interview partners effectively predicts communication accommodation and social status predictions. *Journal of Personality and Social Psychology* 70: 1231-1240.

Gregory, S. W., S. Webster and G. Huang. 1993. Voice pitch and amplitude convergence as a metric of quality in dyadic interviews. *Language Communication* 13: 195-217.

Halle, P. A., C. T. Best and A. Levitt. 1999. Phonetic vs. phonological influences on French listeners' perception of American English approximants. *Journal of Phonetics* 27: 281-306.

Hauser, M. D. 1996. *The Evolution of Communication*. Cambridge, MA: MIT Press.

Henderson, J. B. and B. H. Repp. 1982. Is a stop consonant released when followed by another stop consonant? *Phonetica* 39: 71-82.

Hintzman, D. L. 1986. "Schema abstraction" in a multiple-trace memory model. *Psychological Review* 93: 411-428.

Honorof, D. N., J. Weihing and C. A. Fowler. 2011. Articulatory events are imitated under rapid shadowing. *Journal of Phonetics* 39: 18-38.

Horner, V. and A. Whiten. 2005. Casual knowledge and imitation/emulation switching in chimpanzees (*Pan troglodytes*) and children (*Homo sapiens*). *Animal Cognition* 8: 164-181.

Jamieson, D. G. and D. E. Moroson. 1986. Training non-native speech contrasts in adults: Acquisition of the English /ð/ - /Ɵ/ contrast by francophones. *Perception and Psychophysics* 40: 205-215.

Jassem, W. 1974. *Fonetyka Języka Angielskiego*. 4th edition. Warszawa: PWN.

Jenkins, J. J. and G. H. Yeni-Komshian. 1995. Cross-language speech perception: Perspective and promise. In W. Strange (ed.) *Speech Perception and Linguistic Experience: Issues in Cross-Language Speech Research*. York: York Press: 463-479.

Johnson, K. 1997. Speech perception without speaker normalization: An exemplar model. In K. Johnson and J. W. Mullennix (eds) *Talker Variability in Speech Processing*. San Diego: Academic Press: 145-166.

Jones, D. 1956. *An Outline of English Phonetics*. Cambridge: Heffer and Sons.

Kondaurova, M. V. and A. L. Francis. 2008. The relationship between native allophonic experience with vowel duration and perception of the English tense/lax vowel contrast by Spanish and Russian listeners. *Journal of the Acoustical Society of America* 124: 3959-3971.

Kopczyński, A. 1977. *Polish and American English Consonant Phonemes*. Warszawa: PWN.

Kuhl, P. and A. N. Meltzoff. 1996. Infant vocalizations in response to speech: Vocal imitation and developmental change. *Journal of the Acoustical Society of America* 100: 2425-2438.

Ladefoged, P. 1975. *A Course in Phonetics*. Orlando: Harcourt Brace.

Mańkowska, A., M. Nowacka and M. Kłoczowska. 2009. *"How Much Wood Would a Woodchuck Chuck?": English Pronunciation Practice Book*. Kraków: Wydawnictwo WSE.

McAllister, R., F. E. Flege and T. Piske. 2002. The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics* 30: 229-258.

McClaskey, C. L., D. B. Pisoni and T. D. Carrell. 1983. Transfer of training of a new linguistic contrast in voicing. *Perception and Psychophysics* 34: 323-330.

Meltzoff, A. and M. Moore. 1999. Persons and representation: Why infant imitation is important for theories of human development. In J. Nadel and G. Butterworth (eds) *Imitation in Infancy*. Cambridge: Cambridge University Press: 9-35.

Mitterer, H. and M. Ernestus. 2008. The link between speech perception and production is phonological and abstract: Evidence form the shadowing task. *Cognition* 109: 168-173.

Munro, M. J., T. M. Derwing, J. E. Flege. 1999. Canadians in Alabama: A perceptual study of dialect acquisition in adults. *Journal of Phonetics* 27: 385-403.

Nagell, K., K. Olguin and M. Tomasello. 1993. Processes of social learning in tool use in chimpanzees (*Pan troglodytes*) and human children (*Homo sapiens*). *Journal of Comparative Psychology*. 107: 174-186.

Namy, L. L., L. C. Nygaard and D. Sauerteig. 2002. Gender differences in vocal accommodation: The role of perception. *Journal and Language and Social Psychology* 21: 422-432.

Natale, M. 1975. Convergence of mean vocal intensity in dyadic communication as a function of social desirability. *Journal of Personality and Social Psychology* 32: 790-804.

Nielsen, K. 2011. Specificity and abstractness of VOT imitation. *Journal of Phonetics* 39: 132-14.

Nosofsky, R. M. 1986. Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General* 115: 39-57.

Pardo, J. S. 2006. On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America* 119: 2382-2392.

Pardo, J. S. 2010. Expressing oneself in conversational interaction. In E. Morsella (ed) *Expressing Oneself/Expressing One's Self. Communication, Cognition, Language, and Identity*. New York: Psychology Press: 183-196.

Pardo, J. S., I. Cajori Jay and R. M. Krauss. 2010. Conversational role influences speech imitation. *Attention, Perception and Psychophysics* 72: 2254-2264.

Pardo, J. S., R. Gibbons, A. Suppes and R. M. Krauss. 2012. Phonetic convergence in college roommates. *Journal of Phonetics* 40: 190-197.

Pekkola, J., M. Laasonen, V. Ojanen, T. Autti, L. P. Jaaskelainen, T. Kujala et al. 2006. Perception of matching and conflicting audiovisual speech in dyslexic and fluent readers: An fMRI study at 3T. *NeuroImage* 29: 797-807.

Pierrehumbert, J. B. 2006. The next toolkit. *Journal of Phonetics* 34: 516-530.

Pisoni, D. B., R. N. Aslin, A. J. Perey and B. L. Hennessy. 1982. Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants. *Journal of Experimental Psychology: Human Perception and Performance* 8: 297-314.

Pruitt, P. and J. van Santen. 2006. Training the perception of Hindi dental and retroflex stops by native speakers of English and Japanese. *Journal of the Acoustical Society of America* 119: 1684-1696.

Pulvermuller, F., M. Huss, F. Kherif, F. Moscoso del Prado Martin, O. Hauk and Y. Shtyrov. 2006. Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences of the United States of America* 103: 7865-7870.

Randolph, M. 1989. *Syllable-based Constraints on Properties of English Sounds*. Doctoral Dissertation, Massachusetts Institute of Technology.

Rizzolatti, G. and M. A. Arbib. 1998. Language within our grasp. *Trends in Neurosciences* 21: 188-194.

Rizzolatti, G. and L. Craighero. 2004. The mirror-neuron system. *Annual Review of Neuroscience* 21: 169-192.

Rizzolatti, G., L. Fogassi and V. Gallese. 2001. Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience* 2: 661-670.

Roach, P. *English Phonetics and Phonology*. 3rd edition. Cambridge: Cambridge University Press.

Rojczyk, A. 2008. Release burst in Polish homorganic stop geminates. *Linguistica Silesiana* 29: 75-86.

Rojczyk, A. 2012a. Spontaneous phonetic imitation of L2 vowels in a rapid shadowing task. Paper presented at Pronunciation in Second Language Learning and Teaching 2012. Vancouver Canada, August 24-25 2012.

Rojczyk, A. 2012b. Phonetic and phonological mode in second-language speech: VOT imitation. Paper presented at EUROSLA 2012, Poznań Poland, September 5-8 2012.

Schwartz, J-L., A. Basirat, L. Ménard and M. Sato. 2012. The Perception-for-Action-Control Theory (PACT): A perceptuo-motor theory of speech perception. *Journal of Neurolinguistics* 25: 336-354.

Shepard, C. A., H. Giles and B. A. Le Poire. 2001. Communication accommodation theory. In W. P. Robinson and H. Giles (eds) *The New Handbook of Language and Social Psychology*. Chichester: John Wiley & Sons.

Shockley, K., L. Sabadini and C. A. Fowler. 2004. Imitation in shadowing words. *Perception and Psychophysics* 66: 422-429.

Skipper, J. I., V. Van Wassenhove, H. C. Nusbaum and S. L. Small. 2007. Hearing lips and seeing voices: How cortical areas supporting speech production mediate audiovisual speech perception. *Cerebral Cortex* 17: 2387-2399.

Sobkowiak, W. 2001. *English Phonetics for Poles*. Poznań: Wydawnictwo Poznańskie.

Street, R. L. and H. Giles. 1982. Speech accommodation theory: A social cognitive approach to language and speech behavior. In M. Roloff and C. Berger (eds) *Social Cognition and Communication*. Beverly Hills: Sage.

Triandis, H. C. and L. M. Triandis. 1960. Race, social class, religion, and nationality as determinants of social distance. *Journal of Abnormal and Social Psychology* 61: 110-118.

Trudgill, P. 1986. *Dialects in Contact*. New York: Blackwell Publishing.

Whiten, A. and D. M. Custance. 1996. Studies of imitation in chimpanzees and children. In C. M. Heyes and B. G. Galef (eds) *Social Learning in Animals: The Roots of Culture*. San Diego: Academic Press: 291-318.

Wierzchowska, B. 1980. *Fonetyka i Fonologia Języka Polskiego*. Wrocław: PAN.

Wilson, S. M. and M. Iacoboni. 2006. Neural responses to non-native phonemes varying in producibility: Evidence for the sensimotor nature of speech perception. *NeuroImage* 33: 316-325.

Wilson, S. M., A. P. Saygin, M. I. Sereno and M. Iacoboni. 2004. Listening to speech activates motor areas involved in speech production. *Nature Neuroscience* 7: 701-702.